Journal of Heredity 2010:101(6):789–793
doi:10.1093/jhered/esq069
Advance Access publication June 18, 2010

Cost-Effective, Species-Specific Microsatellite Development for the Endangered Dwarf Bulrush (*Typha minima*) Using Next-Generation Sequencing Technology

Daniela Csencsics, Sabine Brodbeck, and Rolf Holderegger

Research Unit Ecological Genetics and Evolution, Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland.

Address correspondence to D. Csencsics at the address above, or e-mail: daniela.csencsics@wsl.ch.

The dwarf bulrush (*Typha minima* Funck ex Hoppe) is an endangered pioneer plant species of riparian flood plains. In Switzerland, only 3 natural populations remain, but reintroductions are planned. To identify suitable source populations for reintroductions, we developed 17 polymorphic microsatellite markers with perfect repeats using the 454 pyrosequencing technique and tested them on 20 individuals with low-cost M13 labeling. We detected 2 to 7 alleles per locus and found expected and observed heterozygosities of 0.05–0.76 and 0.07–1, respectively. The whole process was finished in less than 6 weeks and cost approximately USD 5000. Due to low costs and reduced expenditure of time, the use of next-generation sequencing techniques for microsatellite development represent a powerful tool for population genetic studies in nonmodel species, as we show in this first application of the approach to a plant species of conservation importance.

Key words: *conservation biology, M13 labeling, riparian flood plains,* Rohrbachia minima, *454 sequencing*

The dwarf bulrush (*Typha minima* Funck ex Hoppe, synonym *Rohrbachia minima* [Funck ex Hoppe] Mavrodiev) is a diploid pioneer plant species of Alpine riparian flood plains. It is listed as endangered in Switzerland (Moser et al. 2002), where only 3 natural populations survived (Käsermann 1999). In other European countries, the species is either extirpated (Germany) or restricted to small populations (Müller 1991, 2007; Csencsics et al. 2008). Larger populations only survive in France (Werner 2001). Several countries undertook specific conservation efforts, such as Austria (Müller 2007) or Switzerland (Camenisch 2000), and in the course of river restoration projects, *T. minima* is stated as a target species to be reintroduced where possible. Although the species' habitat requirements are well known, only 2 studies have dealt with genetic aspects of *T. minima* so far (Galeuchet et al. 2002; Till-Bottraud et al. 2010). In the face of the species' severe decline in the 20th century (Endress 1975), a genetic survey of the remaining populations seems necessary to estimate regional genetic differentiation and to identify suitable source populations for reintroductions.

Microsatellites, highly variable DNA sequences of tandem repeats of 1–6 nucleotides with codominant inheritance, have become the markers of choice for a variety of applications in conservation and population genetics or evolution (Goldstein and Schlötterer 1999). Although the application of microsatellites is simple and robust, their identification and development represent significant challenges. Traditionally, there are 2 approaches to develop polymerase chain reaction (PCR) primers: constructing a genomic library and developing microsatellites de novo or testing known microsatellite primers already developed for related species. With the rise of next generation sequencing, large numbers of sequences can be obtained at greatly reduced prices and in less time than with traditional Sanger sequencing. The identification of microsatellites from genomic DNA using next generation sequencing is relatively new, and only a few studies reported this approach in detail (Allentoft et al. 2008; Abdelkrim et al. 2009; Castoe et al. 2009; Lee et al. 2009; Santana et al. 2009; Tangphatsornruang et al. 2009; Vanpé et al. 2009). In this paper, we describe the development of species-specific microsatellite loci for the endangered dwarf bulrush using 454 sequencing after Abdelkrim et al. (2009). Our goal was to further explore the benefit of 454 sequencing for microsatellite development in combination with cost-effective M13 labeling of candidate primers in microsatellite evaluation.

Materials and Methods

Sample Collection and DNA Extraction

For microsatellite development and testing, we sampled fresh leaf material of *T. minima* from 3 different regions: Lech Valley (12 individuals) and Lake Konstanz (5 individuals) in Austria and Rhine Valley (3 individuals) in Switzerland. Samples were dried in Silica gel, and genomic DNA was extracted with the DNeasy Plant mini kit (QIAGEN) according to the manufacturer's protocol.

454 Sequencing

One DNA sample from Lake Konstanz was then subject to shotgun sequencing using a Roche 454 Genome Sequencer FLX with the Titanium Sequencing kit XLR 70 at Microsynth AG (Balgach, Switzerland). We ordered one-sixteenth run. Proceedings followed the manufacturer's protocol. The technique is described in detail in Margulies et al. (2005).

Data Analysis and Primer Selection

We obtained 76 692 reads in FASTA format with an average read length of 341.3 bp and a total amount of 26 173 179 bases. These sequences should be randomly distributed over the nuclear and potentially chloroplast genome. As we were only interested in nuclear microsatellite markers, loci have to be heterozygous at least in 1 tested sample. We directly screened all the unassembled sequences for perfect di-, tri-, and tetranucleotides using MSATCOMMANDER version 0.8.2 (Faircloth 2008). We chose the option "Design Primers", in which case the software searches for microsatellite repeats and identifies possible primer annealing sites in one single step. We did not use any primer tagging option. PRIMER3 (Rozen and Skaletsky 2000) is used by MSATCOMMANDER for primer design, which means primers are designed to meet the following criteria: amplification products within a size range of 100–500 bp, optimal melting temperature of 60.0 °C (range 57.0 °C–62.0 °C), optimal GC content of 50%, possession of at least one 1 bp GC clamp, low levels of self- or pair-complementarity and maximum end stability (ΔG) of 8.0 (Faircloth 2008).

As our goal was to develop a set of about 15–20 markers, we conducted several screenings with different search criteria (i. e. different minimum repeat lengths for di-, tri-, and tetranucleotides) in order to extract a total of 50–100 sequences with microsatellite repeats and primer annealing sites meeting the strict PRIMER3 criteria. We detected 307 sequences with tandem repeats: 43 di-, 163 tri-, and 101 tetranucleotide microsatellites with at least 8, 10, and 6 repeats, respectively. One hundred of these sequences contained suitable primer annealing sites (Table 1).

We tested these 100 sequences using the CLC SEQUENCE VIEWER 6.0.2 to detect possible alignments of sequences that represented the same target microsatellite locus with identical PCR primer annealing sites but 454 pyrosequencing reads of different length. We found 2 pairs of identical target microsatellite loci, thus reducing the number of potential microsatellite markers to 98, of which 30 were chosen for testing.

Primer Testing

We tested 20 samples of *T. minima* for microsatellite amplification and performed PCRs as follows. The PCR volume of 10 μl contained approximately 1 ng of genomic DNA, 1× Master Mix (Multiplex PCR Kit; QIAGEN), 0.01 μM of forward primer, and 0.15 μM of each reverse and universal FAM-labeled M13 primers (Schuelke 2000). Amplification was carried out separately for each locus on Veriti thermocyclers (Applied Biosystems) with initial denaturing at 94 °C for 15 min, followed by 30 cycles of 94 °C for 30 s, 57 °C for 90 s, and 72 °C for 1 min, followed by 8 cycles of 94 °C for 30 s, 53 °C for 90 s, and 72 °C for 1 min with a final extension of 72 °C for 30 min. PCR products were run with GeneScan-500 LIZ as size standard on a 3130 Genetic Analyzer (Applied Biosystems), and electropherograms were analyzed using GENEMAPPER version 3.7 (Applied Biosystems).

Data Analysis

For all polymorphic loci providing clear electropherograms, we estimated the overall number of alleles and expected and observed heterozygosities using CERVUS 3.0 (Kalinowski et al. 2007) and conducted exact Hardy–Weinberg tests using GENEPOP (Raymond and Rousset 1995; Rousset 2008).

Results

We tested 30 potential microsatellite markers (Table 1). We successfully amplified all tested loci, but 8 primer pairs produced unclear patterns in the electropherograms. These loci would need PCR optimization for reliable scoring of the allele sizes. Of the remaining 22 loci producing scorable peaks, 1 was homozygous and 4 were monomorphic across the 20 samples tested, whereas 17 loci were polymorphic and heterozygous in at least 1 sample. Table 2 presents the characteristics of the 17 polymorphic loci. Across the 20 tested samples, we detected 15 different multilocus genotypes. All identical multilocus genotypes (i.e., a clone) stemmed from the same populations and were collected between 0.5 and 3 m apart from each other. We found between 2 and 7 alleles (mean = 3.67) per locus. Expected and observed heterozygosities ranged from 0.05 to 0.76 (mean = 0.51) and from 0.05 to 1 (mean = 0.38), respectively.

The whole process from DNA extraction to tested primers was finished in less than 6 weeks. In Table 3, we summarize the workflow steps, time, and costs required for the identification of SSRs using 454 sequencing. Once the primers are tested by M13 labeling, a further step would be to multiplex several primers in a single PCR, using conventional fluorescent labeling of microsatellite primers.

Table 1 MSATCOMMANDER/PRIMER3 search criteria and results: minimum number of repeats (minimum no. of repeats), number of detected sequences with tandem repeats (no. of sequences), number of sequences with optimal primer sites (no. of primer pairs), and number of loci selected for testing (no. of tested loci)

	Minimum no. of repeats	No. of sequences	No. of primer pairs	No. of tested loci
Dinucleotide	8	43	5	5
Trinucleotide	10	163	60	23
Tetranucleotide	6	101	33	2
Total		307	98	30

Discussion

We successfully developed 17 polymorphic microsatellite loci for the endangered dwarf bulrush. 6 loci significantly deviated from Hardy–Weinberg equilibrium, most probably due to Wahlund effect as we pooled samples from different regions for analysis. Other possible reasons for the observed deviation from Hardy–Weinberg equilibrium could be presence of null alleles or large allele dropout (Wattier et al. 1998).

Compared with the traditional development of microsatellites from a genomic library as provided by many commercial suppliers, costs were reduced to about a third, and we also reached a reduction in time expenditure. Although we could finish the whole process in less than 6 weeks with the potential to be faster, traditional methods to develop microsatellite primers de novo could easily take several months. We further reduced costs by using

Table 2 Characteristics of 17 polymorphic microsatellite loci in *Typha minima*, tested with 20 samples and excluding ramets of the same genet, which resulted in 15 different multilocus genotypes, including locus name, GenBank accession number repeat motif, primer sequences, total number of alleles, allele size (bp), expected heterozygosity (H_e), observed heterozygosity (H_o), and significance of deviation from Hardy–Weinberg equilibrium at *P*-levels 0.05 (*), 0.01 (**), 0.001 (***), and Bonferroni corrected

Locus	GenBank accession number	Repeat motif	Primer sequence (5′–3′)	Total number of alleles	Allele size (bp)	H_e	H_o
Tmin01	GU936577	(AG)	F: CTTCTTCTCGTGTCCACCG R: TGCAGTACGGCCTCATCG	5	292–308	0.60	0.40
Tmin02	GU936578	(AG)	F: TCCGCCATTAGAGCAACCC R: GCCCTCAGAGTAGCAACC	2	309–311	0.05	0.05
Tmin03	GU936579	(ATT)	F: GGCCTCTGGAGAGTAAGGC R: ACCGAGCCAAGATATCGAAG	6	227–248	0.52	0.40*
Tmin04	GU936580	(ACT)	F: TGGAGACCTTCGAACAGCG R: CCACGACGTAGCAAATTCCC	7	198–246	0.74	1.00**
Tmin05	GU936581	(CTT)	F: TCCGGTTTCAAATTGCTCCAC R: TCTACGCCTCGACTTGGAC	2	238–247	0.44	0.33
Tmin06	GU936582	(ATT)	F: AACAGAGCTGGTCTAGGCG R: GGTCTGATACGGTCCTATGCC	3	307–325	0.57	0.47
Tmin07	GU936583	(CTT)	F: CACTGCCAATCGCCAACTC R: TGATCGATGTTCATTGCTGCC	3	293–308	0.40	0.13*
Tmin08	GU936584	(CTT)	F: ACCGCCATTACTATGCTCC R: TCATGTCTCGAAGGCGTGG	3	202–232	0.50	0.07***
Tmin09	GU936585	(ATT)	F: AGCCTTACGGTTTGGTCAG R: GCATCGGTTATCTGAAGCACC	4	199–208	0.60	0.33
Tmin10	GU936586	(TAA)	F: GGTCGAGCCACTATTGTTTGG R: CAACCTTGTCATGCTCCCG	3	414–423	0.60	0.07***
Tmin11	GU936587	(ATA)	F: TGTTTATATGCTTGACGCACTTC R: CAGACTTGAACCTCCGAATGAC	3	296–302	0.52	0.53
Tmin12	GU936588	(TAT)	F: AACCCGCTACCTTCTCGTG R: GTGCGAGCTTGAGCCATC	7	185–215	0.76	0.50**
Tmin13	GU936589	(TTC)	F: GGGTCTGCCTGGAAATAGC R: CCATCTTGGCCAGCAATGAG	2	182–185	0.23	0.53
Tmin14	GU936590	(TGC)	F: GAAGAAGCAAATACGTGGGAC R: AGATTCTGCCGCCTAGACC	2	277–280	0.44	0.33
Tmin15	GU936591	(TGG)	F: GAAGAAGGCGTCGTAGATTTG R: GTGTGCACCTTGACCTACC	3	170–185	0.53	0.57
Tmin16	GU936593	(TAA)	F: GTTTCGAGCCTTCCAACTCG R: GATAAGTACGGCCCACAACG	3	227–236	0.52	0.57
Tmin17	GU936594	(AACT)	F: TTCAAACCATCCGCCGTTC R: CATCACCCGAAGGCCAAAC	2	194–198	0.50	0.20

Table 3 Overview of workflow steps, required time, and costs for microsatellite marker development using 454 sequencing and M13 labeling during primer evaluation

Workflow steps	Required time	Costs (in USD, including value-added tax)
Extraction of genomic DNA in required concentration	<1 day	15
Library preparation and 454 sequencing at commercial supplier	3 weeks (supplier dependent)	3500
Data analysis: tandem repeat search and primer design	<1 day	—
Primer testing (M13 labeled)	2 weeks (facility dependent)	1500 (consumables, facility dependent)
Total	Less than 6 weeks	≈5000

cheap M13 labeling during primer evaluation (Schuelke 2000).

Apart from time and cost-effectiveness, the enhancement of flexibility is a main advantage of microsatellite development using the 454 technique. The large amount of sequences allows a stepwise selection of loci. In a first step, one can define high-quality standards for tandem repeat number and primer design and only accept loci meeting all criteria. In our case, we obtained enough primers with this first approach, searching for di-, tri-, and tetranucleotides. In a second step, the required tandem repeat number can be reduced, which will increase the number of eligible loci. Microsatellite loci typically consist of 5–40 repeats (Selkoe and Toonen 2006), and several studies describe microsatellite markers with a minimum of 4–5 repeats (Culley 2005; Brzyski 2010; Pazouki et al. 2010). Researchers can also search for penta- and hexanucleotides, which may provide useful markers as well (e.g., Culley et al. 2008). Should there be need for even more loci, then primer design criteria can be relaxed, for example, through accepting a broader range of primer melting temperatures, or compound, imperfect tandem repeats can be accepted. We succeeded in developing all microsatellite markers with primers having optimal annealing temperatures around 60 °C. This enabled the use of consistent PCR conditions across loci and facilitates PCR multiplex composition.

The approach we used here for microsatellite development in a threatened plant species for the first time should allow researchers to use larger number of microsatellites in many studies of nonmodel organisms, as the development of about 20 microsatellite primer pairs will no longer be a serious obstacle due to costs and labor.

Funding

Swiss Federal Office for the Environment BAFU (00.0294.PZ/I472-0912).

Acknowledgments

We thank N. Müller and B. Koch of the Naturpark Tiroler Lech and P. Werner for various support and information on *T. minima* and D. Zulliger and two anonymous referees for detailed and very constructive comments on the manuscript.

References

Abdelkrim J, Robertson BC, Stanton J-AL, Gemmell NJ. 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. BioTechniques. 46:185–192.

Allentoft ME, Schuster SC, Holdaway RN, Hale ML, McLay E, Oskam C, Gilbert TP, Spencer P, Willerslev E, Bunce M. 2008. Identification of microsatellites from an extinct moa species using high-throughput (454) sequence data. BioTechniques. 46:195–200.

Brzyski JR. 2010. Isolation and characterization of microsatellite markers in the rare clonal plant, *Spiraea virginiana* (Rosaceae). Am J Bot. 97:E20–E22, doi: 10.3732/ajb.1000008.

Camenisch M. 2000. Erhaltung und Wiederansiedlung des Kleinen Rohrkolbens (*Typha minima* Hoppe) in Graubünden. Chur, Switzerland: Amt für Landschaftspflege und Naturschutz des Kantons Graubünden.

Castoe TA, Poole AW, Gu W, De Koning J, Daza JM, Smith EN, Pollock DD. 2009. Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. Mol Ecol Resour. 10:341–347.

Csencsics D, Galeuchet D, Keel A, Lambelet C, Müller N, Werner P, Holderegger R. 2008. Der Kleine Rohrkolben—Bedrohter Bewohner eines seltenen Lebensraums. Merkblatt für die Praxis WSL. 43:1–8.

Culley T. 2005. Characterization of newly developed microsatellite loci in the stemmed yellow violet, Viola pubescens (Violaceae). Mol Ecol Notes. 5:882–884.

Culley T, Weller SG, Sakai AK, Putnam KA. 2008. Characterization of microsatellite loci in the Hawaiian endemic shrub *Schiedea adamantis* (Caryophyllaceae) and amplification in related species and genera. Mol Ecol Res. 8:1081–1084.

Endress PK. 1975. Verbreitungsrückgang von *Myricaria germanica* Desv. und *Typha minima* Hoppe auf der Alpennordseite Graubündens. Vierteljschr. der Naturforschenden Gesellschaft Zürich. 120:1–14.

Faircloth BC. 2008. MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour. 8:92–94.

Galeuchet D, Holderegger R, Rutishauser R, Schneller J. 2002. Isozyme diversity and reproduction of *Typha minima* populations on the upper River Rhine. Aquat Bot. 74:19–32.

Goldstein DB, Schlötterer C. 1999. Microsatellites: evolution and applications. Oxford: Oxford University Press.

Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol. 16:1099–1106.

Käsermann C. 1999. Typha minima. In: Merkblätter Artenschutz—Blütenpflanzen und Farne. Berne, Switzerland: BUWAL. p. 284–285.

Lee RM, Thimmapuram J, Thinglum KA, Gong G, Hernandez AG, Wright CL, Kim RW, Mikel MA, Tranel PJ. 2009. Sampling the waterhemp (*Amaranthus tuberculatus*) genome using pyrosequencing technology. Weed Sci. 57:465–469.

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen ZT, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 437:376–380.

Moser DM, Gygax A, Bäumler B, Wyler N, Palese R. 2002. Rote Liste der gefährdeten Arten der Schweiz: farn- und Blütenpflanzen. Berne, Switzerland: BAFU.

Müller N. 1991. Verbreitung, Vergesellschaftung und Rückgang des Zwergrohrkolbens (*Typha minima* Hoppe). Hoppea. 50:19.

Müller N. 2007. Zur Wiederansiedlung des Zwergrohrkolbens (*Typha minima* Hoppe) in den Alpen—eine Zielart alpiner Flusslandschaften. Natur in Tirol. 13:180–193.

Pazouki L, Mardi M, Shanjani PS, Hagidimitriou M, Pirseyedi SM, Naghavi MR, Avanzato D, Vendramin E, Kafkas S, Ghareyazie B, et al. 2010. Genetic diversity and relationships among *Pistacia* species and cultivars. Conserv Genet. 11:311–318.

Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered. 86:248–249.

Rousset F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol Ecol Resour. 8:103–106.

Rozen S, Skaletsky HJ. 2000. Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. Bioinformatics methods and protocols: methods in molecular biology. Totowa, NJ: Humana Press. p. 365–386.

Santana QC, Coetzee MPA, Steenkamp ET, Mlonyeni OX, Hammond GNA, Wingfield MJ, Wingfield BD. 2009. Microsatellite discovery by deep sequencing of enriched genomic libraries. BioTechniques. 46:217–223.

Schuelke M. 2000. An economic method for the fluorescent labelling of PCR fragments. Nat Biotechnol. 18:233–234.

Selkoe KA, Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Let. 9:615–629.

Tangphatsornruang S, Somta P, Uthaipaisanwong P, Chanprasert J, Sangsrakru D, Seehalak W, Sommanas W, Tragoonrung S, Srinives P. 2009. Characterization of microsatellites and gene contents from genome shotgun sequences of mungbean (*Vigna radiata* (L.) Wilczek). BMC Plant Biol. 9. doi:10.1186/1471-2229-9-137.

Till-Bottraud I, Poncet BN, Rioux D, Girel J. 2010. Spatial structure and clonal distribution of genotypes in the rare *Typha minima* Hoppe (Typhaceae) along a river system. Botanica Helv. doi: 10.1007/s00035 -010-0069-x.

Vanpé C, Buschiazzo E, Abdelkrim J, Morrow C, Nicol SC, Gemmell NJ. 2009. Development of microsatellite markers for the short-beaked echidna using three different approaches. Aust J Zool. 57:219–224.

Wattier R, Engel CR, Saumitou-Laprade P, Valero M. 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). Mol Ecol. 7:1569–1573.

Werner P. 2001. Observations sur la distribution des Orthoptères des zones alluviales dans les Alpes occidentales et sur leur valeur d'indicateurs pour la revitalisation des grandes rivières. Bull Rom Entomol. 19:27–46.

Received March 1, 2010; Revised May 10, 2010; Accepted May 20, 2010

Corresponding Editor: David Wagner